SIGNIFICANCE

Physical methods are poised to transform drug discovery and chemical biology via quantitative, predictive design of small molecules.

Unfortunately, these methods still have severe limitations and much work is needed to expand their domain of applicability, but retrospective tests are unsuitable for this.

To truly advance these methods, we need a series of blind challenges focused on pushing the limits of predictive techniques.

SAMPL has been vital to fulfill this role; here we propose to continue and extend SAMPL via collection of a set of carefully selected experimental data in order to drive further improvements in modeling.

This work is necessary in order to advance modeling to the point where it can reliably guide drug discovery efforts, reducing time consuming and costly trial and error.

INNOVATION

We will discuss how SAMPL has previously driven innovation in the field.

Innovation here will include developing new, high-throughput experiments for studying protein ligand binding (Aim 3).

In Aim 4, we will not only run SAMPL community challenges, but also perform our own reference calculations with the latest techniques, testing their accuracy and using these to assess the current state-of-the-art.

Innovation here will also involve careful assessment of how to compare methods and analyze their relative performance in a statistically sound way.

APPROACH

Aim 1: Generate new data for "simple" SAMPL blind challenges on physical property prediction. We will develop new solution-phase datasets for druglike small molecules. These data can test critical aspects of small molecule modeling (including accounting for interactions and treatment of protonation/tautomeric state) and improve our ability to predict physical properties relevant to drug discovery in new regions of chemical space. We will initially focus on distribution between organic phases and on pKa?s and their modulation by solvent environment, using these data to drive improvements in the modeling of ligand interactions.

Aim 2: Measure binding of novel host-guest complexes for introductory ligand binding challenges. We will measure new host-guest binding free energies for cucurbiturils and deep-cavity cavitands, yielding further host-guest binding challenges which span between physical property prediction and protein-ligand binding. Host guest systems are some of the simplest cases of molecular recognition, and thus these binding data will drive improvements in modeling of simple binding systems with techniques of relevance to drug discovery.

Cucubituril derivatives for host-guest binding. The Isaacs group has previously participated in the SAMPL challenges and supplied unpublished host-guest binding constants [**?**, **?**, **?**]. Our participation was quite stimulating for us and influenced our investigation of the biomedical applications of acyclic CB[n]-type receptors (a.k.a Calabadions). Cucurbit[n]uril receptors are particularly well suited for the SAMPL challenges because they exhibit: 1) high binding constants toward suitable guests in water (routinely μM to nM; occasionally pM to fM) [1–7], 2) high selectivities between structurally related guests which translate into large $\Delta\Delta G$ values [8], 3) low molecular weights (1000-2000 amu) which allows high levels of theory to be used, and 4) limited conformational degrees of freedom. Herein, we propose to continue to participate in the next three SAMPL challenges during the proposed five year funding period by resynthesizing previously published CB[n]-type receptors of increasing complexity, measure Ka values and determine host-guest stoichiometry and geometry toward biologically relevant guests which will allow the computational chemists to push the boundaries of the free-energy prediction of receptor?ligand complexes. Figure **??** shows the chemical structures of three hosts – Me4CB[8] [9], glycoluril hexamer [10], and acyclic CB[n]-type receptors [11–16] which span the range from preorganized macrocyclic host to uncharged acyclic but preorganized host to highly charged acyclic host.

SAMPL6. For this challenge we propose to measure K_a and ΔH values, stoichiometry, and geometry for the interaction of Me4CB[8] (nicely water soluble CB[8] derivative) toward 15 guests (chosen from top selling drugs, Table CB1) by either direct or competition isothermal titration calorimetry (ITC), UV/Vis or fluorescence indicator displacement assay, or NMR competition experiments which we are very experienced with [1, 2, 17, 18]. Our selection of Me_4CB[8] and top 100 drugs was based on a desire to increase the level of complexity of the computational challenge by: 1) changing host flexibility (e.g. Me_4CB[8] can exhibit ellipsoidal deformation) [9], 2) by allowing the possibility of binary or ternary (e.g. 1:1 and/or 1:2 host:guest) complexes [19–21], 3) using

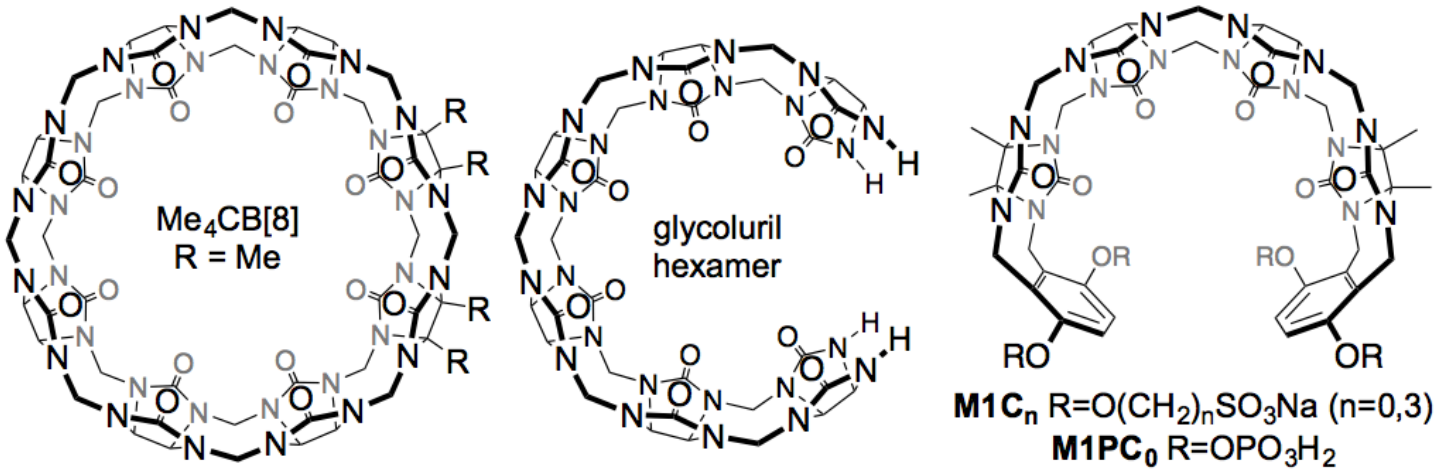

Figure 1: **Structures of Me_4CB[8], glycouril hexamer, and acyclic CB[n]-type receptors.**

drugs with several potential binding epitopes to include sampling issues. Host:guest stoichiometry and geometry (e.g. which binding epitope is complexed) will be addressed by ITC "n" values, Job plots monitored by UV/Vis or NMR [22], and by 1H NMR complexation induced changes in chemical shifts [23]. All three sets of studies will be conducted in phosphate buffered saline (pH 7.4 with physiological salt) which introduces further complexity due to competitive interaction between the C=O portals of CB[n]-type receptors and metal ions via ion-dipole interactions which reduces the observed Ka values [24].

drug	**features**
memantine	adamantane; 1:1
saxagliptin	adamantane; 1:1
premarin	steroid
pancuronium	steroid
varenicline	1:1 vs 1:2
valsartan	pKa 4.37
omeprazole	pKa 4.77
ranolazine	pKa 7.17; epitopes
pradaxa	pKa 3.87; epitopes
nilotinib	epitopes; pKa 6.3
sensipar	epitopes; folding
vyvance	diamine; epitopes; folding
minocycline	tetracyclin; amino aniline

Table 1: Selected drugs as guests

SAMPL8. We propose to study host:guest complexes of glycoluril hexamer toward the 15 drugs (Figure CB1). We select glycoluril hexamer for this challenge because it: 1) increases the conformational dynamics of the host, and 2) influences the number and energy of solvating (and unusually coordinated) water molecules that are implicated in the observed high binding constants for CB[n]-guest complexes [7, 25]. Furthermore, in selecting the drugs, we have chosen several that have pKa values in the 3.8 to 7.4 range. Similar to biomolecular host-guest systems, CB[n]-type receptors are well known to induce pKa shifts (up to 4 pKa units) of complexed guests [26–28], and the ability of computation to replicate and predict such shifts and their impact on Ka are of high significance.

SAMPL10 We will focus on acyclic CB[n]-type receptors (e.g. M1C_3, M1C_0, and M1PC_0 that contain anionic solubilizing groups attached via different linker lengths. As in SAMPL2, these acyclic CB[n]-type receptor introduces conformational complexity and influences the free energy of the solvating H_2O molecules in the free host. Moreover, the presence of 4 anionic groups in close proximity to the cavity are expected to have a significant influence on the balance between ion-dipole interactions and the solvation of the free host.

Gibb deep cavity cavitands for host-guest studies. During SAMPL4 [29] and SAMPL5 [30] we focused on two hosts: the octa-acid 1 (R = H) and another octa-acid derivative with four methyl groups positioned at the portal of the binding pocket (1, R = Me). These studies used Isothermal Titration Calorimetry (ITC) to measure the thermodynamics of respectively: 1 (R = H) complexing a range of carboxylate guests, and the binding of carboxylate and trimethylammonium guests to both hosts (1, H = H and Me). In both cases NMR was also used in a confirmatory role for free energy data. SAMPL5 emphasized that slight differences in the shape of the hydrophobic pocket of the host can in the case of some guests have a profound affect on affinity.

To continue with this work we will expand on the range of hosts by including 2 and 3 in our ITC studies. Like cavitand 1, host 2 is an octa-acid derivative. However, the four benzoate groups are relocated from the extreme exterior in the case of 1, to the rim of the binding pocket in 2. We surmise that this will have a direct effect on the binding of charged guests, but more subtly, an indirect effect on guest complexation via changes to the solvation of the empty host. Octa-trimethylammonuim cavitand ("positand" 3) has the same overall architecture as host 1, but

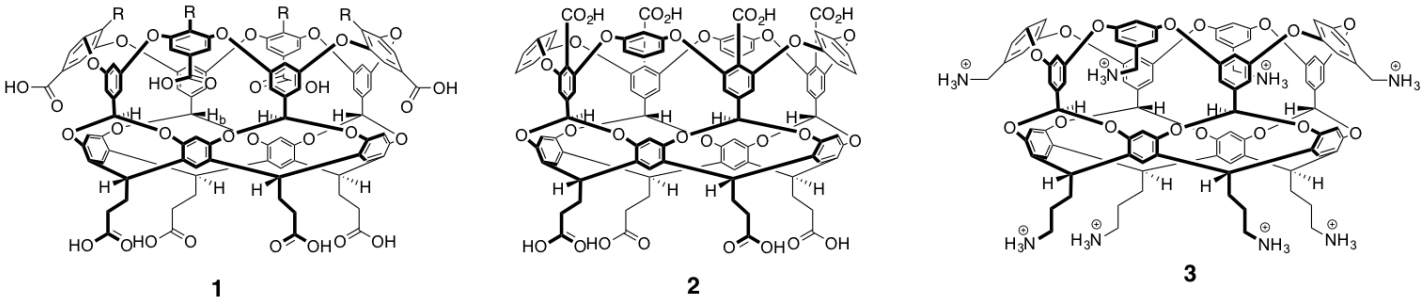

Figure 2: **Gibb deep cavity cavitands for SAMPL6-10.**

inverts the charges on the water solubilizing exterior coat. It is not entirely clear at this juncture if this switch in groups relatively remote from the pocket can directly affect guest complexation. However related (unpublished) results suggest that it can.

Guests for the five proposed ITC studies will be obtained from commercial sources and will be selected on both the limitations of force-fields available to the computationalists and new data as it is gathered.

SAMPL6 will involve hosts 1 and 2 with a set of five, previously uninvestigated guests. The principal aim for our group will be to examine how the location of the carboxylate affects guest binding. Building on what the computational chemists learn from this study, SAMPL7 will compare hosts 1 and 3 with a different set of five guests. We anticipate that because of the relative remoteness of the charged groups in these two hosts, the effects of switching charges will be subtler than the differences between 1 and 2. SAMPL8 will switch gears and consider the effects salts have on guest binding. Thus, we will compare the effects of NaCl and NaI on the complexation of five guests to 1. We have previously shown that iodide has a weak affinity for the binding pocket of 1, whilst sodium ions have an affinity for the outer carboxylates [31], and this will be made clear to the participants. We will follow on from this with SAMPL9 looking at the affects of these same two salts on the complexation of five guests to 3. We have not yet quantified the complexation of this salt to host 3, but expect the iodide to have affinity for both the pocket and the positively charged solubilizing groups. Finally, for SAMPL10 we will consider the effects of co-solvents on the binding of five guests to 1 and 2. This is for us an entirely new area, but we expect binding to be weaker because of co-solvent affinity for the binding pocket leading to competition; there will be an apparent weakening of the hydrophobic effect. However, the precise nature of this weakening phenomenon are unclear.

Aim 3. Generate biologically relevant advanced model systems for protein-ligand binding challenges. We will identify suitable biological protein-ligand model systems (difficult but tractable in order to push the limits of physical techniques) then measure binding and develop these for blind challenges. This will include binding studies on human serum albumin and bromodomains or aspartyl proteases; initial binding data will be expanded by the selection of additional ligands or the creation of mutations in the protein that modulate binding.

Aim 4. Coordinate, run, and analyze blind challenges to advance modeling of binding. The data collected in Aims 1-3 will drive annual SAMPL blind challenges, allowing the field to test the latest methods and force fields to assess progress, compare them against one another head-to-head, and perform sensitivity analysis to learn how much different factors (protonation state, tautomer selection, solvent model, force field, sampling method, etc.) affect predictive power. Results will then feed back into improved treatment of these factors for subsequent challenges, driving regular cycles of application, learning, and advancement.

TIMELINE

COLLABORATION MANAGEMENT PLAN

OUTLOOK

Bibliography and References Cited

[1] Cao, L., Šekutor, M., Zavalij, P. Y., Mlinarić-Majerski, K., Glaser, R., and Isaacs, L.: Cucurbit[7]uril·Guest Pair with an Attomolar Dissociation Constant. Angew. Chem. Int. Ed. 53(4): 988–993, January 2014.

[2] Liu, S., Ruspic, C., Mukhopadhyay, P., Chakrabarti, S., Zavalij, P. Y., and Isaacs, L.: The Cucurbit[n]uril Family: Prime Components for Self-Sorting Systems. Journal of the American Chemical Society. 127(45): 15959–15967, November 2005.

[3] Mock, W. L. and Shih, N. Y.: Structure and selectivity in host-guest complexes of cucurbituril. The Journal of Organic Chemistry. 51(23): 4440–4446, November 1986.

[4] Assaf, K. I. and Nau, W. M.: Cucurbiturils: From synthesis to high-affinity binding and catalysis. Chem Soc Rev. 44(2): 394–418, January 2015.

[5] Moghaddam, S., Yang, C., Rekharsky, M., Ko, Y. H., Kim, K., Inoue, Y., and Gilson, M. K.: New Ultrahigh Affinity Host-Guest Complexes of Cucurbit[7]uril with Bicyclo[2.2.2]octane and Adamantane Guests: Thermodynamic Analysis and Evaluation of M2 Affinity Calculations. Journal of the American Chemical Society. 133(10): 3570–3581, March 2011.

[6] Shetty, D., Khedkar, J. K., Park, K. M., and Kim, K.: Can we beat the biotin–avidin pair?: cucurbit[7]uril-based ultrahigh affinity host–guest complexes and their applications. Chem. Soc. Rev. 44(23): 8747–8761, 2015.

[7] Biedermann, F., Uzunova, V. D., Scherman, O. A., Nau, W. M., and De Simone, A.: Release of High-Energy Water as an Essential Driving Force for the High-Affinity Binding of Cucurbit[n]urils. J. Am. Chem. Soc. 134(37): 15318–15323, September 2012.

[8] Isaacs, L.: Stimuli Responsive Systems Constructed Using Cucurbit[n]uril-Type Molecular Containers. Acc. Chem. Res. 47(7): 2052–2062, July 2014.

[9] Vinciguerra, B., Zavalij, P. Y., and Isaacs, L.: Synthesis and Recognition Properties of Cucurbit[8]uril Derivatives. Org. Lett. 17(20): 5068–5071, October 2015.

[10] Lucas, D., Minami, T., Iannuzzi, G., Cao, L., Wittenberg, J. B., Anzenbacher, P., and Isaacs, L.: Templated Synthesis of Glycoluril Hexamer and Monofunctionalized Cucurbit[6]uril Derivatives. J. Am. Chem. Soc. 133(44): 17966–17976, November 2011.

[11] Ma, D., Zhang, B., Hoffmann, U., Sundrup, M. G., Eikermann, M., and Isaacs, L.: Acyclic Cucurbit[n]uril-Type Molecular Containers Bind Neuromuscular Blocking Agents In Vitro and Reverse Neuromuscular Block In Vivo. Angew. Chem. Int. Ed. 51(45): 11358–11362, November 2012.

[12] Ma, D., Hettiarachchi, G., Nguyen, D., Zhang, B., Wittenberg, J. B., Zavalij, P. Y., Briken, V., and Isaacs, L.: Acyclic cucurbit[n]uril molecular containers enhance the solubility and bioactivity of poorly soluble pharmaceuticals. Nat Chem. 4(6): 503–510, June 2012.

[13] Zhang, B. and Isaacs, L.: Acyclic Cucurbit[n]uril-type Molecular Containers: Influence of Aromatic Walls on their Function as Solubilizing Excipients for Insoluble Drugs. J. Med. Chem. 57(22): 9554–9563, November 2014.

[14] Gilberg, L., Zhang, B., Zavalij, P. Y., Sindelar, V., and Isaacs, L.: Acyclic cucurbit[n]uril-type molecular containers: Influence of glycoluril oligomer length on their function as solubilizing agents. Org. Biomol. Chem. 13(13): 4041–4050, 2015.

[15] Sigwalt, D., Moncelet, D., Falcinelli, S., Mandadapu, V., Zavalij, P. Y., Day, A., Briken, V., and Isaacs, L.: Acyclic Cucurbit[n]uril-Type Molecular Containers: Influence of Linker Length on Their Function as Solubilizing Agents. ChemMedChem. 11(9): 980–989, May 2016.

[16] Zhang, B., Zavalij, P. Y., and Isaacs, L.: Acyclic CB[n]-type molecular containers: Effect of solubilizing group on their function as solubilizing excipients. Org. Biomol. Chem. 12(15): 2413–2422, 2014.

[17] Ma, D., Zavalij, P. Y., and Isaacs, L.: Acyclic Cucurbit[n]uril Congeners Are High Affinity Hosts. J. Org. Chem. 75(14): 4786–4795, July 2010.

[18] She, N., Moncelet, D., Gilberg, L., Lu, X., Sindelar, V., Briken, V., and Isaacs, L.: Glycoluril-Derived Molecular Clips are Potent and Selective Receptors for Cationic Dyes in Water. Chem. Eur. J. pp n/a–n/a, August 2016.

[19] Ko, Y. H., Kim, E., Hwang, I., and Kim, K.: Supramolecular assemblies built with host-stabilized charge-transfer interactions. Chem. Commun. (13): 1305–1315, 2007.

[20] Barrow, S. J., Kasera, S., Rowland, M. J., del Barrio, J., and Scherman, O. A.: Cucurbituril-Based Molecular Recognition. Chem. Rev. 115(22): 12320–12406, November 2015.

[21] Urbach, A. R. and Ramalingam, V.: Molecular Recognition of Amino Acids, Peptides, and Proteins by Cucurbit[n]uril Receptors. Isr. J. Chem. 51(5-6): 664–678, May 2011.

[22] Connors, K. A.: Binding Constants. New York, NY, John Wiley & Sons, 1987.

[23] Masson, E., Ling, X., Joseph, R., Kyeremeh-Mensah, L., and Lu, X.: Cucurbituril chemistry: A tale of supramolecular success. RSC Adv. 2(4): 1213–1247, 2012.

[24] Márquez, C., Hudgins, R. R., and Nau, W. M.: Mechanism of Host-Guest Complexation by Cucurbituril. J. Am. Chem. Soc. 126(18): 5806–5816, May 2004.
[25] Biedermann, F., Nau, W. M., and Schneider, H.-J.: The Hydrophobic Effect Revisited—Studies with Supramolecular Complexes Imply High-Energy Water as a Noncovalent Driving Force. Angew. Chem. Int. Ed. 53(42): 11158–11171, October 2014.
[26] 'il Saleh, N., Koner, A., and Nau, W.: Activation and Stabilization of Drugs by Supramolecular pKa Shifts: Drug-Delivery Applications Tailored for Cucurbiturils. Angewandte Chemie. 120(29): 5478–5481, July 2008.
[27] Nau, W. M., Florea, M., and Assaf, K. I.: Deep Inside Cucurbiturils: Physical Properties and Volumes of their Inner Cavity Determine the Hydrophobic Driving Force for Host–Guest Complexation. Isr. J. Chem. 51(5-6): 559–577, May 2011.
[28] Ghosh, I. and Nau, W. M.: The strategic use of supramolecular pKa shifts to enhance the bioavailability of drugs. Advanced Drug Delivery Reviews. 64(9): 764–783, June 2012.
[29] Gibb, C. L. D. and Gibb, B. C.: Binding of cyclic carboxylates to octa-acid deep-cavity cavitand. J Comput Aided Mol Des. 28(4): 319–325, November 2013.
[30] Sullivan, M. R., Sokkalingam, P., Nguyen, T., Donahue, J. P., and Gibb, B. C.: Binding of carboxylate and trimethylammonium salts to octa-acid and TEMOA deep-cavity cavitands. J Comput Aided Mol Des. pp 1–8, July 2016.
[31] Carnegie, R. S., Gibb, C. L. D., and Gibb, B. C.: Anion Complexation and The Hofmeister Effect. Angew. Chem. 126(43): 11682–11684, October 2014.